JAPAN PATENT OFFICE

This is to certify that the annexed is a true copy of the following application as filed with this office.

Date of Application :	January 27, 2004
Application Number :	Patent Application No. 2004-018344 [JP2004-018344]
Applicant(s) :	RIKEN MEDICAL & BIOLOGICAL LABORATORIES CO., LTD.

January 14, 2005

Hiroshi OGAWA

Commissioner, Patent Office

Certification No. 2004-312265

VERIFICATION OF TRANSLATION

I, Makoto AIHARA, Patent Attorney, of SIKs & Co., 8th Floor, Kyobashi-Nisshoku Bldg., 8-7, Kyobashi 1-chome, Chuo-ku, Tokyo 104-0031 JAPAN hereby declare that I am the translator of the certified official copy of the documents in respect of an application for a patent filed in Japan on January 27, 2004 under Patent Application No. 018344/2004 and that the following is a true and correct translation to the best of my knowledge and belief.

Dated: September 17, 2010

Makoto AIHARA

[Title]	Application for Patent
[Reference Number]	A31700A
[Filing Date]	January 27, 2004
[Address]	To the Commissioner of the Patent Office
[Inventor]	
[Domicile or Residence]	c/o RIKEN, 2-1, Hirosawa, Wako-shi, Saitama
[Name]	Atsushi MIYAWAKI
[Inventor]	
[Domicile or Residence]	c/o RIKEN, 2-1, Hirosawa, Wako-shi, Saitama
[Name]	Takako KOGURE
[Inventor]	
[Domicile or Residence]	c/o MEDICAL & BIOLOGICAL LABORATORIES CO., LTD., INALAB., 1063-103, Aza Ohara, Oaza Terasawaoka, Ina-shi, Nagano
[Name]	Satoshi KARASAWA
[Applicant for Patent]	
[Identification Number]	503359821
[Name]	RIKEN
[Applicant for Patent]	
[Identification Number]	390004097
[Name]	MEDICAL & BIOLOGICAL LABORATORIES CO., LTD.
[Administrator]	
[Identification Number]	110000109
[Name]	SIKs & Co.
[Representative]	Masazumi IMAMURA
[Fee]	
[Number of register of payment]	170347
[Amount]	21,000
[List of Attached Documents]	
[Document's Name]	Claims 1
[Document's Name]	Specification 1
[Document's Name]	Drawings 1
[Document's Name]	Abstract 1
[Number of Comprehensive Power of Attorney]	0316216

[Name of Document] CLAIMS

[Claim 1] A chromoprotein described in the following (a) or (b):

(a) a protein having the amino acid sequence shown in SEQ ID NO: 1; or

(b) a protein, which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 1, and which has light-absorbing properties.

[Claim 2] A fluorescent protein described in the following (a) or (b):

(a) a protein having the amino acid sequence shown in SEQ ID NO: 3; or

(b) a protein, which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 3, and which has fluorescence properties.

[Claim 3] A fluorescent protein described in the following (a) or (b):

(a) a protein having the amino acid sequence shown in SEQ ID NO: 5 or 7; or

(b) a protein, which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 5 or 7, which has fluorescence properties, and which has a stokes shift of 100 nm or greater.

[Claim 4] DNA encoding a chromoprotein described in the following (a) or (b):

(a) a protein having the amino acid sequence shown in SEQ ID NO: 1; or

(b) a protein, which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 1, and which has light-absorbing properties.

[Claim 5] DNA encoding a fluorescent protein described in the following (a) or (b):

(a) a protein having the amino acid sequence shown in SEQ ID NO: 3; or

(b) a protein, which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 3, and which has fluorescence properties.

[Claim 6] DNA encoding a fluorescent protein described in the following (a) or (b):

(a) a protein having the amino acid sequence shown in SEQ ID NO: 5 or 7; or

(b) a protein, which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 5 or 7, which has fluorescence properties, and which has a stokes shift of 100 nm or greater.

[Claim 7] DNA described in the following (a) or (b):

(a) DNA having the nucleotide sequence shown in SEQ ID NO: 2; or

(b) DNA, which has a nucleotide sequence comprising a deletion, substitution, and/or addition of one or several nucleotides with respect to the nucleotide sequence shown in SEQ ID NO: 2, and which has a nucleotide sequence encoding a protein that has light-absorbing properties.

[Claim 8] DNA described in the following (a) or (b):

(a) DNA having the nucleotide sequence shown in SEQ ID NO: 4; or

(b) DNA, which has a nucleotide sequence comprising a deletion, substitution, and/or addition of one or several nucleotides with respect to the nucleotide sequence shown in SEQ ID NO: 4, and which has a nucleotide sequence encoding a protein that has fluorescence properties.

[Claim 9] DNA described in the following (a) or (b):

(a) DNA having the nucleotide sequence shown in SEQ ID NO: 6 or 8; or

(b) DNA, which has a nucleotide sequence comprising a deletion, substitution, and/or addition of one or several nucleotides with respect to the nucleotide sequence shown in SEQ ID NO: 6 or 8, and which has a nucleotide sequence encoding a protein that has fluorescence properties and has a stokes shift of 100 nm or greater.

[Claim 10] A recombinant vector having the DNA according to any one of claims 4 to 9.

[Claim 11] A transformant having the DNA according to any one of claims 4 to 9 or the

recombinant vector according to claim 10.

[Claim 12] A fusion protein, which consists of the protein according to any one of claims 1 to 3 and another protein.

[Claim 13] The fusion protein according to claim 12, wherein another protein is a protein that localizes in a cell.

[Claim 14] The fusion protein according to claim 12 or 13, wherein another protein is a protein specific to a cell organella.

[Claim 15] A method for analyzing the localization or dynamics of a protein in a cell, which is characterized in that the fusion protein according to any one of claims 12 to 14 is allowed to express in the cell.

[Claim 16] A reagent kit, which comprises: the fluorescent protein of any one of claims 1 to 3; the DNA of any one of claims 4 to 9; the recombinant vector of claim 10; the transformant of claim 11; or the fusion protein of any of claims 12 to 14.

[Name of Document] SPECIFICATION

[Title of Invention] FLUORESCENT PROTEIN

[TECHNICAL FIELD]

[0001]

The present invention relates to a novel chromoprotein and fluorescent protein. More specifically, the present invention relates to a novel chromoprotein and fluorescent protein derived from Montipora. sp, and use thereof.

[BACKGROUND ART]

[0002]

Green fluorescent protein (GFP) derived from Aequorea victoria, a jellyfish, has many purposes in biological systems. Recently, various GFP mutants have been produced based on the random mutagenesis and semi-rational mutagenesis, wherein a color is changed, a folding property is improved, luminance is enhanced, or pH sensitivity is modified. Fluorescent proteins such as GFP are fused with other proteins by gene recombinant technique, and monitoring of the expression and transportation of the fusion proteins is carried out.

[0003]

One of the most commonly used types of GFP mutant is Yellow fluorescent protein (YFP). Among Aequorea-derived GFP mutants, YFP exhibits the fluorescence with the longest wavelength. The values ε and Φ of the majority of YEPs are 60,000 to 100,000 $M^{-1}cm^{-1}$ and 0.6 to 0.8, respectively (Tsien, R. Y. (1998). Ann. Rev. Biochem. 67, 509-544). These values are comparable to those of the general fluorescent group (fluorescein, rhodamine, etc.). Accordingly, improvement of the absolute luminance of YFP is nearly approaching its limit.

[0004]

In addition, cyan fluorescent protein (CFP) is another example of the GFP mutant. Of this type of protein, ECFP (enhanced cyan fluorescent protein) has been known.

Moreover, red fluorescent protein (RFP) has been isolated from sea anemone (Discoma sp.). Of this type of protein, DasRed has been known. Thus, 4 types of fluorescent proteins, that are, green fluorescent protein, yellow fluorescent protein, cyan fluorescent protein, and red fluorescent protein, have successively been developed. The range of the spectrum has significantly been expanded.

[0005]

Several GFP homologs derived from Aequorea have a large stokes shift (the difference between an excitation peak value and a fluorescence peak value) (GFPuv and sapphire). These GFP homologs obtain green fluorescence as a result of excitation with UV light at 380 nm. However, the use of such UV light having toxicity is not suitable for observation in organisms. No red fluorescent proteins have a large stokes shift. Under the current circumstances, either excitation or fluorescence must be sacrificed in fluorescence observation.

[0006]

[Non-patent document 1] Tsien, R. Y. (1998). Ann. Rev. Biochem. 67, 509-544

[DISCLOSURE OF THE INVENTION]

[Object to be solved by the Invention]

[0007]

When compared with a low molecular weight fluorescent substance, a fluorescent protein has broad excitation and fluorescence spectra. Many fluorescent proteins have overlapped portions between such excitation and fluorescence spectra. Thus, it is extremely difficult to excite at an excitation peak value and then to observe at a fluorescence peak value. It is an object of the present invention to provide a fluorescent protein which is able to solve the aforementioned problem. That is to say, it is an object of the present invention to provide a red or orange fluorescent protein, which is characterized in that the difference (stokes shift) between an excitation peak value (wavelength of maximum absorption) and a fluorescence peak value (wavelength of

maximum fluorescence) is greatened, so that the maximum fluorescence can be obtained by the maximum excitation.

[Means for solving the object]

[0008]

As a result of intensive studies directed towards achieving the aforementioned objects, the present inventors have attempted to isolate a gene encoding a novel chromoprotein using Montipora sp. as a material, so as to obtain a chromoprotein COCP. Subsequently, the present inventors have substituted histidine at position 94 of the COCP protein with asparagine, asparagine at position 142 with serine, asparagine at position 157 with aspartic acid, lysine at position 201 with arginine, and phenylalanine at position 206 with serine, so as to produce a fluorescent protein COCP-FL having fluorescence properties. COCP-FL has an excitation peak at 560 nm, and because of this excitation, the peak of the fluorescence spectrum was obtained at 600 nm. Further, the present inventors have substituted serine at position 61 of the aforementioned COCP-FL with phenylalanine, isoleucine at position 92 with threonine, valine at position 123 with threonine, phenylalanine at position 158 with tyrosine, valine at position 191 with isoleucine, and serine at position 213 with alanine, so as to produce a protein keima 616, which has fluorescence properties that are different from those of COCP-FL. Such keima 616 has an excitation peak at 440 nm, and because of this excitation, the peak of the fluorescence spectrum was obtained at 616 nm. Its stokes shift was 176 nm, which was an extremely large value. Still further, the present inventors have substituted phenylalanine at position 61 of keima 616 with methionine, and glutamine at position 62 with cysteine, so as to produce a fluorescent protein keima 570. Such keima 570 has an excitation peak at 440 nm as with keima 616, and because of this excitation, it has a fluorescence peak at 570 nm. Its stokes shift was 130 nm, which was a large value. The present invention has been completed based on these findings.

[0009]

The present invention provides a chromoprotein described in the following (a) or (b):

(a) a protein having the amino acid sequence shown in SEQ ID NO: 1; or

(b) a protein, which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 1, and which has light-absorbing properties.

[0010]

Another aspect of the present invention provides a fluorescent protein described in the following (a) or (b):

(a) a protein having the amino acid sequence shown in SEQ ID NO: 3; or

(b) a protein, which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 3, and which has fluorescence properties.

[0011]

Further another aspect of the present invention provides a fluorescent protein described in the following (a) or (b):

(a) a protein having the amino acid sequence shown in SEQ ID NO: 5 or 7; or

(b) a protein, which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 5 or 7, which has fluorescence properties, and which has a stokes shift of 100 nm or greater.

[0012]

Further another aspect of the present invention provides DNA encoding a chromoprotein described in the following (a) or (b):

(a) a protein having the amino acid sequence shown in SEQ ID NO: 1; or

(b) a protein, which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence

shown in SEQ ID NO: 1, and which has light-absorbing properties.

[0013]

Further another aspect of the present invention provides DNA encoding a fluorescent protein described in the following (a) or (b):

(a) a protein having the amino acid sequence shown in SEQ ID NO: 3; or

(b) a protein, which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 3, and which has fluorescence properties.

[0014]

Further another aspect of the present invention provides DNA encoding a fluorescent protein described in the following (a) or (b):

(a) a protein having the amino acid sequence shown in SEQ ID NO: 5 or 7; or

(b) a protein, which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 5 or 7, which has fluorescence properties, and which has a stokes shift of 100 nm or greater.

[0015]

Further another aspect of the present invention provides DNA described in the following (a) or (b):

(a) DNA having the nucleotide sequence shown in SEQ ID NO: 2; or

(b) DNA, which has a nucleotide sequence comprising a deletion, substitution, and/or addition of one or several nucleotides with respect to the nucleotide sequence shown in SEQ ID NO: 2, and which has a nucleotide sequence encoding a protein that has light-absorbing properties.

[0016]

Further another aspect of the present invention provides DNA described in the following (a) or (b):

(a) DNA having the nucleotide sequence shown in SEQ ID NO: 4; or

(b) DNA, which has a nucleotide sequence comprising a deletion, substitution, and/or addition of one or several nucleotides with respect to the nucleotide sequence shown in SEQ ID NO: 4, and which has a nucleotide sequence encoding a protein that has fluorescence properties.

[0017]

Further another aspect of the present invention provides DNA described in the following (a) or (b):

(a) DNA having the nucleotide sequence shown in SEQ ID NO: 6 or 8; or

(b) DNA, which has a nucleotide sequence comprising a deletion, substitution, and/or addition of one or several nucleotides with respect to the nucleotide sequence shown in SEQ ID NO: 6 or 8, and which has a nucleotide sequence encoding a protein that has fluorescence properties and has a stokes shift of 100 nm or greater.

[0018]

Further another aspect of the present invention provides a recombinant vector having the DNA according to the present invention as mentioned above.

Further another aspect of the present invention provides a transformant having the DNA or the recombinant vector according to the present invention as mentioned above.

[0019]

Further another aspect of the present invention provides a fusion protein, which consists of the protein according to the present invention as mentioned above and another protein. Preferably, said another protein is a protein that localizes in a cell. More preferably, said another protein is a protein specific to a cell organella.

[0020]

Further another aspect of the present invention provides a method for analyzing the localization or dynamics of a protein in a cell, which is characterized in that the

fusion protein according to the present invention as mentioned above is allowed to express in the cell.

Further another aspect of the present invention provides a reagent kit, which comprises: the fluorescent protein, the DNA, the recombinant vector, the transformant or the fusion protein according to the present invention as mentioned above.

[Effect of the Invention]

[0021]

The fluorescent proteins of the present invention emit red and orange fluorescence, and the excitation peak thereof is 440 nm (blue). The conventional red fluorescent proteins (DsRed and HcRed) have a stokes shift (the difference between an excitation peak value and a fluorescence peak value) between 20 and 30 nm. In contrast, the red fluorescent protein of the present invention has a stokes shift of 176 nm, and the orange fluorescent protein of the present invention has a stokes shift of 130 nm. Thus, the fluorescent proteins of the present invention have extremely large values. Accordingly, the fluorescent protein of the present invention is characterized in that the maximum fluorescence can be obtained by the maximum excitation. Moreover, since the excitation peak is at 440 nm, in the simultaneous excitation staining with a cyan fluorescent protein (CFP) or a green fluorescent protein (GFP), it becomes possible to extremely effectively obtain the fluorescence of both proteins. Furthermore, the excitation peak of the conventional red fluorescent proteins is between 560 nm and 590 nm. In contrast, the fluorescent protein of the present invention has an excitation peak at 440 nm. Thus, by changing excitation light, it makes possible to stain the present fluorescent protein, simultaneously with the conventional red fluorescent protein.

[BEST MODE FOR CARRYING OUT THE INVENTION]

[0022]

The embodiments of the present invention will be described in detail below.

<u>(1) Proteins of the present invention</u>

The proteins of the present invention are: a protein having the amino acid sequence shown in SEQ ID NO: 1, 3, 5 or 7; and a protein, which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 1, 3, 5 or 7, and which has absorption properties or fluorescence properties. The stokes shifts (the difference between the wavelength of maximum absorption and the wavelength of maximum fluorescence) of the proteins having the amino acid sequence shown in SEQ ID NO: 5 or 7, are 176 nm and 130 nm, respectively. The stokes shifts of the proteins, which have an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 5 or 7, and which have fluorescence properties, are 100 nm or greater, and more preferably 120 nm or greater.

[0023]

The proteins of the present invention are characterized in that they have the following properties:

(1) COCP (the amino acid sequence thereof is shown in SEQ ID NO: 1, and the nucleotide sequence thereof is shown in SEQ ID NO: 2)

Wavelength of maximum excitation (wavelength of maximum absorption): 576 nm

Molar absorption coefficient at 576 nm: 64,000

pH sensitivity: absent

[0024]

(2) COCP-FL (the amino acid sequence thereof is shown in SEQ ID NO: 3, and the nucleotide sequence thereof is shown in SEQ ID NO: 4)

Wavelength of maximum excitation (wavelength of maximum absorption): 560 nm

Wavelength of maximum fluorescence: 600 nm

[0025]

(3) keima 616 (the amino acid sequence thereof is shown in SEQ ID NO: 5, and the

nucleotide sequence thereof is shown in SEQ ID NO: 6)

Wavelength of maximum excitation (wavelength of maximum absorption): 440 nm

Wavelength of maximum fluorescence: 616 nm

pH sensitivity: fluorescence intensity is stable between pH 7.5 and 10

[0026]

(4) keima 570 (the amino acid sequence thereof is shown in SEQ ID NO: 7, and the nucleotide sequence thereof is shown in SEQ ID NO: 8)

Wavelength of maximum excitation (wavelength of maximum absorption): 440 nm

Wavelength of maximum fluorescence: 570 nm

pH sensitivity: fluorescence intensity is stable between pH 7.5 and 10

[0027]

In the examples of the present specification, DNA encoding the protein of the present invention was cloned from Montipora sp. used as a starting material. Montipora sp. is a certain type of coral, which belongs to Acropora, Scleractinia, Hexacorallia, Anthozoa, Cnidaria. It often forms an aggregated or coated colony. It is to be noted that the protein of the present invention can also be obtained from coral emitting fluorescence other than Montipora sp. in some cases, and that such a protein is also included in the scope of the present invention.

[0028]

The scope of "one or several" in the phrase "an amino acid sequence comprising a deletion, substitution and/or addition of one or several amino acids" used herein is not particularly limited. For example, it means 1 to 20, preferably 1 to 10, more preferably 1 to 7, further preferably 1 to 5, and particularly preferably 1 to 3.

[0029]

The term "protein having light-absorbing properties" is used to mean in the present specification to mean a protein having properties capable of absorbing light with a certain wavelength. The light-absorbing properties of a "protein, which has an amino

acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 1, and which has light-absorbing properties" may be either substantially identical to, or different from those of the protein having the amino acid sequence shown in SEQ ID NO: 1. Such light-absorbing properties can be evaluated based on absorption intensity, excitation wavelength (absorption wavelength), pH sensitivity, etc., for example. Among the proteins of the present invention, chromoproteins, which have light-absorbing properties and do not emit fluorescence, can be used, for example, as (1) an FRET acceptor molecule (energy receptor), or can be used in (2) the development of a system for converting irradiated light energy to energy other than light, or in (3) introduction of a mutation into the amino acid sequence of a protein to modify it, so that it can emit fluorescence.

[0030]

The term "protein having fluorescence properties" is used in the present specification to mean a protein having properties capable of emitting fluorescence as a result of excitation with light having a certain wavelength. The fluorescence properties of the "proteins, which have an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 3, 5 or 7, and which have fluorescence properties" may be either identical to or different from the fluorescence properties of the proteins having the amino acid sequence shown in SEQ ID NO: 3, 5 or 7. Such fluorescence properties can be evaluated based on fluorescence intensity, excitation wavelength, fluorescence wavelength, pH sensitivity, etc., for example.

[0031]

The method of obtaining the fluorescent protein and the chromoproteins of the present invention is not particularly limited. The proteins may be either a protein synthesized by chemosynthesis, or recombinant protein produced by a gene

recombination technique.

[0032]

Where a recombinant protein is produced, it is necessary to obtain DNA encoding the protein. Appropriate primers are designed by using information regarding the amino acid sequence shown in SEQ ID NO: 1, 3, 5 or 7 of the sequence listing of the present specification and the nucleotide sequence shown in SEQ ID NO: 2, 4, 6 or 8 thereof. Using these primers, PCR is carried out by using cDNA library derived from Montipora sp. as a template, so that DNA encoding the protein of the present invention can be obtained. Where a partial fragment of DNA encoding the protein of the present invention is obtained by the above-described PCR, the produced DNA fragments are ligated to one another by a gene recombination technique, so that DNA encoding the desired protein can be obtained. The protein of the present invention can be produced by introducing this DNA into an appropriate expression system. Expression in an expression system will be described later in the present specification.

[0033]

<u>(2) DNA of the present invention</u>

The present invention provides genes encoding the chromo protein or the fluorescent protein of the present invention.

A specific example of DNA encoding the protein of the present invention is

(a) a protein having the amino acid sequence shown in SEQ ID NO: 1, 3, 5 or 7; or

(b) a protein, which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 1, 3, 5 or 7, and which has light-absorbing properties or fluorescence properties.

[0034]

Further specific example of DNA encoding the chromoprotein or fruolescent protein of the present invention is DNA described in the following (a) or (b)

(a) DNA having the nucleotide sequence shown in SEQ ID NO: 2, 4, 6 or 8; or

(b) DNA, which has a nucleotide sequence comprising a deletion, substitution, and/or addition of one or several nucleotides with respect to the nucleotide sequence shown in SEQ ID NO: 2, 4, 6 or 8, and which has a nucleotide sequence encoding a protein that has light-absorbing properties or fluorescence properties.

[0035]

In the term “a nucleotide sequence comprising a deletion, substitution and/or addition of one or several nucleotides” used in the present specification, the range of “one or several” is not particularly limited, but is, for example, from 1 to 50, preferably 1 to 30, more preferably 1 to 20, still more preferably 1 to 10, and particularly preferably 1 to 5.

[0036]

The DNA of the present invention can be synthesized by, for example, the phosphoamidite method, or it can also be produced by polymerase chain reaction (PCR) using specific primers. The DNA of the present invention or its fragment is produced by the method described above in the specification.

[0037]

A method of introducing a desired mutation into a certain nucleic acid sequence is known to a person skilled in the art. For example, known techniques such as a site-directed mutagenesis, PCR using degenerated oligonucleotides, or the exposure of cells containing nucleic acid to mutagens or radioactive rays, are appropriately used, so as to construct DNA having a mutation. Such known techniques are described in, for example, Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY., 1989; and Current Protocols in Molecular Biology, Supplements 1 to 38, John Wiley & Sons (1987-1997).

[0038]

(3) Recombinant vector of the present invention

The DNA of the present invention can be inserted into a suitable vector and used. The type of a vector used in the present invention is not particularly limited. For example, it may be either a vector that can autonomously replicate (e.g., a plasmid, etc.), or vector that is incorporated into the genomes of host cells when it is introduced into the host cells and is then replicated together with the chromosome into which it is incorporated.

[0039]

The vector used in the present invention is preferably an expression vector. In an expression vector, elements necessary for transcription (e.g., a promoter, etc.) are functionally ligated to the DNA of the present invention. The promoter is a DNA sequence which shows a transcriptional activity in host cells, and it is appropriately selected depending on the type of host cells.

[0040]

Examples of a promoter which can operate in bacterial cells may include a *Bacillus stearothermophilus* maltogenic amylase gene promoter, a *Bacillus licheniformis* alpha-amylase gene promoter, a *Bacillus amyloliquefaciens* BAN amylase gene promoter, a *Bacillus subtilis* alkaline protease gene promoter, a *Bacillus pumilus* xylosidase gene promoter, P_R and P_L promoters of phage rhamda, and lac, trp and tac promoters of *Escherichia coli*.

[0041]

Examples of a promoter which can operate in mammalian cells may include an SV40 promoter, an MT-1 (metallothionein gene) promoter, and an adenovirus-2 major late promoter. Examples of a promoter which can operate in insect cells may include a polyhedrin promoter, a P10 promoter, an *Autographa californica* polyhedrosis basic protein promoter, a baculovirus immediate-early gene 1 promoter, and a baculovirus 39K delayed-early gene promoter. Examples of a promoter which can be operate in yeast host cells may include promoters derived from yeast glycolytic genes, an alcohol

dehydrogenase gene promoter, a TPI1 promoter, and an ADH2-4c promoter.

Examples of a promoter which can operate in filamentous cells may include an ADH3 promoter and a tpiA promoter.

[0042]

In addition, an appropriate terminator such as a human growth hormone terminator, or a TPI1 terminator or ADH3 terminator for fungal cells, may be functionally bound to the DNA of the present invention, as necessary. The recombinant vector of the present invention may further have elements such as a polyadenylation signal (e.g., one derived from SV40 or the adenovirus 5E1b region), a transcription enhancer sequence (e.g., an SV40 enhancer), or a translation enhancer sequence (e.g., one encoding the adenovirus VA RNA).

The recombinant vector of the present invention may further comprise a DNA sequence which enables the replication of the recombinant vector in host cells. SV40 replication origin is an example of such a sequence (when the host cells are mammalian cells).

[0043]

The recombinant vector of the present invention may further comprise a selective marker. Examples of such a selective marker may include genes, complements of which are absent from host cells, such as a dihydrofolate reductase (DHFR) gene or a *Shizosaccharomyces pombe* TPI gene, and drug resistant genes such as ampicillin, kanamycin, tetracycline, chloramphenicol, neomycin or hygromycin-resistant genes.

A method for ligating the DNA of the present invention, a promoter and, as desired, a terminator and/or a secretory signal sequence to one another and inserting these items into a suitable vector is known to a person skilled in the art.

[0044]

<u>(4) Transformant of the present invention</u>

A transformant can be produced by introducing the DNA or recombinant vector of the present invention into a suitable host.

Any cell can be used as a host cell into which the DNA or recombinant vector of the present invention is introduced, as long as the DNA construct of the present invention can be expressed therein. Examples of such a cell may include bacteria, yeasts, fungal cells, and higher eukaryotic cells.

[0045]

Examples of bacteria may include Gram-positive bacteria such as Bacillus or Streptomyces, and Gram-negative bacteria such as *Escherichia coli*. These bacteria may be transformed by the protoplast method or other known methods, using competent cells.

Examples of mammalian cells may include HEK 293 cells, HeLa cells, COS cells, BHK cells, CHL cells, and CHO cells. A method of transforming mammalian cells and expressing the introduced DNA sequence in the cells is also known. Examples of such a method may include the electroporation, the calcium phosphate method, and the lipofection method.

[0046]

Examples of yeast cells may include those belonging to Saccharomyces or Shizosaccharomyces. Examples of such cells may include *Saccharomyces cerevisiae* and *Saccharomyces kluyveri*. Examples of a method of introducing a recombinant vector into yeast host cells may include the electroporation, the spheroplast method, and the lithium acetate method.

[0047]

Examples of other fungal cells may include those belonging to *Filamentous fungi* such as Aspergillus, Neurospora, Fusarium or Trichoderma. Where *Filamentous fungi* are used as host cells, transformation can be carried out by incorporating DNA constructs into host chromosomes, so as to obtain recombinant host cells. Incorporation

of DNA constructs into the host chromosomes is carried out by known methods, and such known methods may include homologous recombination and heterologous recombination.

[0048]

Where insect cells are used as host cells, both a vector into which a recombinant gene is introduced and a baculovirus are co-introduced into insect cells, and a recombinant virus is obtained in the culture supernatant of the insect cells. Thereafter, insect cells are infected with the recombinant virus, so as to allow the cells to express proteins (described in, for example, Baculovirus Expression Vectors, A Laboratory Manual; and Current Protocols in Molecular Biology, Bio/Technology, 6, 47 (1988)).

[0049]

The *Autographa californica* nuclear polyhedrosis virus, which is a virus infecting to insects belonging to *Barathra brassicae*, can be used as baculovirus.

Examples of insect cells used herein may include Sf9 and Sf21, which are *Spodoptera frugiperda* ovarian cells [Baculovirus Expression Vectors, A Laboratory Manual, W. H. Freeman & Company, New York, (1992)], and HiFive (manufactured by Invitrogen), which are *Trichoplusia ni* ovarian cells.

Examples of the method of co-introducing both a vector into which a recombinant gene has been introduced and the above baculovirus into insect cells to prepare a recombinant virus may include the calcium phosphate method and the lipofection method.

[0050]

The above transformant is cultured in an appropriate nutritive medium under conditions enabling the introduced DNA construct to be expressed. In order to isolate and purify the protein of the present invention from the culture product of the transformant, common methods of isolating and purifying proteins may be used.

For example, where the protein of the present invention is expressed in a state

dissolved in cells, after completion of the culture, cells are recovered by centrifugal separation, and the recovered cells are suspended in a water type buffer. Thereafter, the cells are disintegrated using an ultrasonic disintegrator or the like, so as to obtain a cell-free extract. A supernatant is obtained by centrifuging the cell-free extract, and then, a purified sample can be obtained from the supernatant by applying, singly or in combination, the following ordinary protein isolation and purification methods: the solvent extraction, the salting-out method using ammonium sulfate or the like, the desalting method, the precipitation method using an organic solvent, the anion exchange chromatography using resins such as diethylaminoethyl (DEAE) sepharose, the cation exchange chromatography using resins such as S-Sepharose FF (manufactured by Pharmacia), the hydrophobic chromatography using resins such as butyl sepharose or phenyl sepharose, the gel filtration method using a molecular sieve, the affinity chromatography, the chromatofocusing method, and the electrophoresis such as isoelectric focusing.

[0051]

(5) Use of the protein of the present invention and a fusion protein comprising the same

The protein of the present invention can be fused with another protein, so as to construct a fusion protein.

A method of obtaining the fusion protein of the present invention is not particularly limited. It may be either a protein synthesized by chemosynthesis, or recombinant protein produced by a gene recombination technique.

Where a recombinant protein is produced, it is necessary to obtain DNA encoding the protein. Appropriate primers are designed using the information regarding the amino acid sequence shown in SEQ ID NO: 1, 3, 5 or 7 of the sequence listing of the present specification and the nucleotide sequence shown in SEQ ID 2, 4, 6 or 8. Using these primers, PCR is carried out using a DNA fragment containing the gene of the protein of the present invention as a template, so as to produce DNA fragments necessary

for construction of the DNA encoding the protein of the present invention. Moreover, DNA fragment encoding a protein to be fused is also obtained in the same above manner.

[0052]

Subsequently, the thus obtained DNA fragments are ligated to one another by a gene recombination technique, so that DNA encoding the desired fusion protein can be obtained. This DNA is then introduced into an appropriate expression system, so that the fusion protein of the present invention can be produced.

[0053]

The fluorescent protein of the present invention has an extremely high utility value as a marker. This is to say, the protein of the present invention is purified as a fusion protein with an amino acid sequence to be tested, and the fusion protein is introduced into cells by methods such as the microinjection. By observing the distribution of the fusion protein over time, targeting activity of the amino acid sequence to be tested can be detected in the cells.

[0054]

The type of another protein (an amino acid sequence to be tested) with which the protein of the present invention is fused is not particularly limited. Preferred examples may include proteins localizing in cells, proteins specific for intracellular organelles, and targeting signals (e.g., a nuclear transport signal, a mitochondrial presequence, etc.). In addition, the protein of the present invention can be expressed in cells and used, as well as being introduced into cells by the microinjection or the like. In this case, a vector into which the DNA encoding the protein of the present invention is inserted in such a way that it can be expressed, is introduced into host cells.

[0055]

Moreover, the protein of the present invention can also be used as a reporter protein to determine promoter activity. This is to say, a vector is constructed such that DNA encoding the fluorescent protein of the present invention is located downstream of

a promoter to be tested, and the vector is then introduced into host cells. By detecting the fluorescence of the fluorescent protein of the present invention which is emitted from the cells, the activity of the promoter to be tested can be determined. The type of a promoter to be tested is not particularly limited, as long as it operates in host cells.

[0056]

A vector used to detect the targeting activity of the above amino acid sequence to be tested or to determine promoter activity is not particularly limited. Examples of a vector preferably used for animal cells may include pNEO (P. Southern, and P. Berg (1982) J. Mol. Appl. Genet. 1: 327), pCAGGS (H. Niwa, K. Yamamura, and J. Miyazaki, Gene 108, 193-200 (1991)), pRc/CMV (manufactured by Invitrogen), and pCDM8 (manufactured by Invitrogen). Examples of a vector preferably used for yeasts may include pRS303, pRS304, pRS305, pRS306, pRS313, pRS314, pRS315, pRS316 (R. S. Sikorski and P. Hieter (1989) Genetics 122: 19-27), pRS423, pRS424, pRS425, pRS426 (T. W. Christianson, R. S. Sikorski, M. Dante, J. H. Shero, and P. Hieter (1992) Gene 110: 119-122).

[0057]

In addition, the type of cells used herein is also not particularly limited. Various types of animal cells such as L cells, BalbC-3T3 cells, NIH3T3 cells, CHO (Chinese hamster ovary) cells, HeLa cells or NRK (normal rat kidney) cells, yeast cells such as *Saccharomyces cerevisiae*, *Escherichia coli* cells, or the like can be used. Vector can be introduced into host cells by common methods such as the calcium phosphate method or the electroporation.

[0058]

The above obtained fusion protein of the present invention wherein the protein of the present invention is fused with another protein (referred to as a protein X) is allowed to be expressed in cells. By monitoring a fluorescence emitted, it becomes possible to analyze the localization or dynamics of the protein X in cells. That is, cells

transformed or transfected with DNA encoding the fusion protein of the present invention are observed with a fluorescence microscope, so that the localization and dynamics of the protein X in the cells can be visualized and thus analyzed.

[0059]

For example, by using a protein specific for an intracellular organella as a protein X, the distribution and movement of a nucleus, a mitochondria, an endoplasmic reticulum, a Golgi body, a secretory vesicle, a peroxisome, etc., can be observed.

Moreover, for example, axis cylinders or dendrites of the nerve cells show an extremely complicated change in strikes in an individual who is under development. Accordingly, fluorescent labeling of these sites enables a dynamic analysis.

[0060]

The fluorescence of the fluorescent protein of the present invention can be detected with a viable cell. Such detection can be carried out using, for example, a fluorescence microscope (Axiophoto Filter Set 09 manufactured by Carl Zeiss) or an image analyzer (Digital Image Analyzer manufactured by ATTO).

The type of a microscope can be appropriately selected depending on purposes. Where frequent observation such as pursuit of a change over time is carried out, an ordinary incident-light fluorescence microscope is preferable. Where observation is carried out while resolution is emphasized, for example, in the case of searching localization in cells specifically, a confocal laser scanning microscope is preferable. In terms of maintenance of the physiological state of cells and prevention from contamination, an inverted microscope is preferable as a microscope system. When an erecting microscope with a high-powered lens is used, a water immersion lens can be used.

[0061]

A filter set can be appropriately selected depending on the fluorescence wavelength of a fluorescent protein. For example, in the case of keima616, the

excitation maximum wavelength (adsorption maximum wavelength) is 440 nm, and the fluorescence maximum wavelength is 616 nm, and therefore a filter having an excitation light between approximately 420 and 460 nm and a fluorescence between approximately 600 and 640 nm is preferably used. In the case of keima570, the excitation maximum wavelength (adsorption maximum wavelength) is 440 nm, and the fluorescence maximum wavelength is 570 nm, and therefore a filter having an excitation light between approximately 420 and 460 nm and a fluorescence between approximately 550 and 590 nm is preferably used.

[0062]

When viable cells are observed over time using a fluorescence microscope, a high sensitive cooled CCD camera is used, since photography is carried out in a short time. In the case of the cooled CCD camera, CCD is cooled to decrease thermal noise, so that a weak fluorescence image can be clearly photographed by exposure in a short time.

[0063]

(6) Kit of the present invention

The present invention provides a kit for analyzing the localization of intracellular components and/or analyzing physiologically active substances, which is characterized in that it comprises at least one selected from the protein, the fusion protein, the DNA, the recombinant vector, or the transformant, which are described in the present specification. The kit of the present invention can be produced from commonly used materials that are known per se, by using common methods.

Reagents such as the fluorescent protein or the DNA are dissolved in an appropriate solvent, so that the reagents can be prepared in a form suitable for conservation. Water, ethanol, various types of buffer solution, etc. can be used as such a solvent.

The present invention will be further described in the following examples.

However, the present invention is not limited by these examples.

[EXAMPLES]

[0064]

Example 1: Isolation of novel chromoprotein gene from stony coral, preparation of novel fluorescent protein, and analysis of properties thereof

(1) Extraction of total RNA

A chromoprotein gene was isolated from coral. Montipora sp. was used as a material. A frozen Montipora sp. was crushed in a mortar, and 7.5 ml of "TRIzol" (GIBCO BRL) was then added to 1 g (wet weight) of the crushed Montipora sp. Thereafter, the obtained mixture was homogenized and then centrifuged at 1,500 x g for 10 minutes. Thereafter, 1.5 ml of chloroform was added to the obtained supernatant, and the mixture was then stirred for 15 seconds. Thereafter, the mixture was left at rest for 3 minutes. The resultant was then centrifuged at 7,500 x g for 15 minutes. Thereafter, 3.75 ml of isopropanol was added to the obtained supernatant, and the mixture was then stirred for 15 seconds, followed by leaving the mixture at rest for 10 minutes. Thereafter, the resultant was centrifuged at 17,000 x g for 10 minutes. The obtained supernatant was discarded, and 6 ml of 70% ethanol was added to the residue, followed by centrifugation at 17,000 x g for 10 minutes. The obtained supernatant was discarded, and the precipitate was then dissolved in 200 μl of DEPC water. Total RNA dissolved in the DEPC water was 100 times diluted, and the values of O.D.260 and O.D.280 were then measured, so as to determine RNA concentration. As a result, 53 μg of total RNA was obtained.

[0065]

(2) Synthesis of first strand cDNA

cDNA (33 μl) was synthesized from 4 μg of the total RNA, using a kit for synthesizing first strand cDNA "Ready To Go" (Amersham Pharmacia).

[0066]

(3) Degenerated PCR

3 μl of the synthesized first strand cDNA (33 μl) was used as a template to carry out PCR. Primers were produced by making comparison among the amino acid sequences of known fluorescent proteins, extracting similar portions, and converting them to nucleotide sequences.

Primers used:

5'-GAAGGRTGYGTCAAYGGRCAY-3' (primer 1) (SEQ ID NO: 9)

5'-ACVGGDCCATYDGVAAGAAARTT-3' (primer 2) (SEQ ID NO: 10)

I represents inosine; R represents A or G; Y represents C or T; V represents A, C, or G; D represents A, G, or T; S represents C or G; H represents A, T, or C

[0067]

Composition of PCR reaction solution:

Template (first strand cDNA)	3 μl
X10 taq buffer	5 μl
2.5 mM dNTPs	4 μl
100 μM primer 1	1 μl
100 μM primer 2	1 μl
Milli-Q	35 μl
taq polymerase (5 U/μl)	1 μl

[0068]

PCR reaction conditions:

94°C x 1 min (PAD)

94°C x 30 sec (denaturation)

52°C x 30 sec (annealing of primers to template)

72°C x 1 min (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 35 times.

72°C x 7 min (final elongation)

4°C (maintenance)

[0069]

Using 1 μl of the amplified product obtained in the first PCR reaction as a template, PCR was carried out again under the same above conditions. A 350-bp fragment was cut out via agarose gel electrophoresis, and it was then purified.

[0070]

(4) Subcloning and sequencing

The purified DNA fragment was ligated to a pT7-blue vector (Novagen). Escherichia coli (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of Escherichia coli, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer. Thereafter, the obtained nucleotide sequence was compared with the nucleotide sequences of other fluorescent protein genes, so as to determine whether the nucleotide sequence of the DNA fragment was derived from a fluorescent protein. With regard to those that were determined to be a part of the fluorescent protein genes, the full-length genes were cloned by the 5′-RACE method and the 3′-RACE method.

[0071]

(5) 5′-RACE method

In order to determine the nucleotide sequence on the 5′-side of the DNA fragment obtained by the degenerated PCR, the 5′-RACE method was applied using 5′-RACE System for Rapid Amplification of cDNA Ends, Version 2.0 (GIBCO BRL). 5 μg of the total RNA prepared in (1) above was used as a template.

The following primers were used in the first amplification of dC-tailed cDNA:

5'-GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG-3' (primer 3) (SEQ ID NO: 11); and

5'-CTCAGGGAATGACTGCTTTACAT-3' (primer 4) (SEQ ID NO: 12)

Herein, I represents inosine.

[0072]

The following primers were used in the second amplification:

5'-GGCCACGCGTCGACTAGTAC-3' (primer5) (SEQ ID NO:13)

5'- GTCTTCAGGGTACTTGGTGA -3' (primer6) (SEQ ID NO:14)

PCR reaction conditions were applied in accordance with protocols attached to the kit.

[0073]

The amplified 350-bp band was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was ligated to a pT7-blue vector (Novagen). Escherichia coli (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of Escherichia coli, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer.

[0074]

(6) 3′-RACE method

The nucleotide sequence on the 3′-side of the DNA fragment obtained by the degenerated PCR was obtained by PCR using a primer prepared based on the information obtained by determination of the nucleotide sequence in (4) above and an oligo dT primer. 3 μl of the first strand cDNA prepared in (2) above was used as a template. The prepared primer was 5'- ATGTAAAGCAGTCATTCCCTGAG -3' (primer7) (SEQ ID NO: 15).

[0075]

Composition of PCR reaction solution:

Template (first strand cDNA)	3 μl
X10 taq buffer	5 μl
2.5 mM dNTPs	4 μl
20 μM primer 7	1 μl

10 μM oligo dT primer	1 μl
Milli-Q	35 μl
Taq polymerase (5 U/μl)	1 μl

[0076]

PCR reaction conditions:

94°C x 1 minute (PAD)

94°C x 30 seconds (denaturation)

52°C x 30 seconds (annealing of primers to template)

72°C x 1 minute (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 30 times.

72°C x 7 minutes (final elongation)

4°C (maintenance)

[0077]

The amplified band with a length of approximately 650 bp was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was ligated to a pT7-blue vector (Novagen). Escherichia coli (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of Escherichia coli, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer.

[0078]

(7) Expression of protein in Escherichia coli

Based on the obtained full-length nucleotide sequence, a primer was produced with a portion corresponding to the N-terminus of the protein. An oligo dT primer was used as a primer corresponding to the C-terminal side thereof. Thereafter, using such primers, PCR was carried out employing the first strand cDNA prepared in (2) above as a template. The obtained full-length amino acid sequence and nucleotide sequence are shown in SEQ ID NOS: 1 and 2 of the sequence listing. This protein having the amino

acid sequence shown in SEQ ID NO: 1 was named COCP.

Primer used:

5'-CCCGGATCCGACCATGGCTACCTTGGTTAAAGA-3' (primer8) (SEQ ID NO: 16)

[0079]

Composition of PCR reaction solution:

Template (first strand cDNA)	3 μl
X10 pyrobest buffer	5 μl
2.5 mM dNTPs	4 μl
100 μM primer 8	1 μl
100 μM oligo dT primer	1 μl
Milli-Q	35 μl
Pyrobest polymerase (5 U/μl)	1 μl

[0080]

PCR reaction conditions:

94°C x 1 minute (PAD)

94°C x 30 seconds (denaturation)

52°C x 30 seconds (annealing of primers to template)

72°C x 1 minute (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 30 times.

72°C x 7 minutes (final elongation)

4°C (maintenance)

[0081]

The amplified band with a length of approximately 800 bp was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was subcloned into the BamHI-EcoRI site of a pRSET vector (Invitrogen), and it was then allowed to express in Escherichia coli (JM109-DE3). The expressed protein was constructed such that His-tag was attached to the N-terminus thereof, and thus it was

purified with Ni-Agarose gel (QIAGEN). Purification was carried out in accordance with the attached protocols. Subsequently, the properties of the purified protein were analyzed.

[0082]

(8) Analysis of light-absorbing properties

Using a solution comprising 20 μM chromoprotein and 50 mM HEPES (pH 7.9), the absorption spectrum was measured. Thereafter, the molar absorption coefficient of the protein was calculated from the value of the peak of the spectrum. In the chromoprotein (COCP) derived from Montipora sp., the absorption peak was observed at 576 nm (Table 1, Figure 1). Also, it was stable at pH4 to 10 (Figure 2).

[0083]

[Table 1:

Properties of chromoprotein (COCP) isolated from Montipora sp.

Table 1

	Maximum absorption	Maximum fluorescence	Molar absorption coefficient	Quantum yield	pH sensitivity	Number of amino acids
COCP	576 nm	-	64000 (576 nm)	-	Absent	221 a.a.
keima 616	440 nm	616 nm	28000 (440 nm)	0.24	Present	222 a.a.

[0084]

(9) Modification from chromoprotein to fluorescent protein

COCP is not a fluorescent protein. However, valine was inserted into the portion between methionine at position 1 of COCP and serine at position 2 thereof, histidine at position 94 was substituted with asparagine, asparagine at position 142 was substituted with serine, asparagine at position 157 was substituted with aspartic acid, lysine at position 202 was substituted with arginine, and phenylalanine at position 206 was substituted with serine, so as to acquire fluorescence properties. This modified fluorescent protein was named as COCP-FL (the amino acid sequence thereof is shown

in SEQ ID NO: 3, and the nucleotide sequence thereof is shown in SEQ ID NO: 4). COCP-FL has an excitation peak at 560 nm. By this excitation, the fluorescence spectrum has a peak at 600 nm.

[0085]

(10) Production of red fluorescent protein with large stokes shift

In COCP-FL, serine at position 62 was substituted with phenylalanine, isoleucine at position 93 was substituted with threonine, valine at position 124 was substituted with threonine, phenylalanine at position 159 was substituted with tyrosine, valine at position 192 was substituted with isoleucine, and serine at position 214 was substituted with alanine, so as to acquire a protein having different fluorescence from that of COCP-FL. This modified fluorescent protein was named as keima 616 (the amino acid sequence thereof is shown in SEQ ID NO: 5, and the nucleotide sequence thereof is shown in SEQ ID NO: 6). keima 616 has an excitation peak at 440 nm. By this excitation, the fluorescence spectrum has a peak at 616 nm (Figure 3, Table 1). Its stokes shift is 176 nm, which is an extremely large value. When compared with the conventional fluorescent protein, this protein is able to have a large excitation wavelength region and a large fluorescence wavelength region, and thus fluorescence can be efficiently measured. In addition, it is also possible to simultaneously measure multiple colors of fluorescence. Using fluorochromes having identical excitation wavelengths, photometry can be conducted with two wavelengths by excitation with a single wavelength such as a laser. In the case of the conventional fluorescent proteins, since proteins having the same excitation spectrum have not existed, such photometry has not been conducted. Using these proteins, a problem regarding deviation in measurement due to difference in excitation can be solved.

[0086]

(11) Production of orange fluorescent protein having large stokes shift

In keima 616, phenylalanine at position 62 was substituted with methionine, and

glutamine at position 63 was substituted with cysteine, so as to obtain a fluorescent protein. This modified fluorescent protein was named as keima 570 (the amino acid sequence thereof is shown in SEQ ID NO: 7, and the nucleotide sequence thereof is shown in SEQ ID NO: 8). As with keima 616, keima 570 also has an excitation peak at 440 nm. By this excitation, the fluorescence spectrum has a peak at 570 nm (Figure 4). Its stokes shift is 130 nm, which is a large value. When compared with the conventional fluorescent protein, this protein is able to have a large excitation wavelength region and a large fluorescence wavelength region, and thus fluorescence can be efficiently measured. In addition, it is also possible to simultaneously measure multiple colors of fluorescence. Using fluorochromes having identical excitation wavelengths, photometry can be conducted with two wavelengths by excitation with a single wavelength such as a laser. In the case of the conventional fluorescent proteins, since proteins having the same excitation spectrum have not existed, such photometry has not been conducted. Using these proteins, a problem regarding deviation in measurement due to difference in excitation can be solved.

[0087]

(12) Measurement of pH sensitivity

The absorption spectra of the proteins (keima 616 and keima 570) were measured in the following 50 mM buffer solutions (Figures 5 and 6):

The pH of each buffer solution is as follows:

pH 4, 5, and 5.5: acetate buffer

pH 6: phosphate buffer

pH 6.6: MOPS buffer

pH 7, 7.5, and 8: HEPES buffer

pH 9 and 10: glycine buffer

The peak value was stable between pH 7.5 and 10 (Figures 5 and 6).

[BRIEF DESCRIPTION OF THE DRAWINGS]

[0088]

[Figure 1] Figure 1 shows the absorption spectrum of COCP.

[Figure 2] Figure 2 shows the results obtained by measuring the pH sensitivity of COCP.

[Figure 3] Figure 3 shows the excitation spectrum and fluorescence spectrum of keima 616.

[Figure 4] Figure 4 shows the excitation spectrum and fluorescence spectrum of keima 570.

[Figure 5] Figure 43 shows the results obtained by measuring the pH sensitivity of keima 616.

[Figure 6] Figure 44 shows the results obtained by measuring the pH sensitivity of keima 570.

[SEQUENCE LISTING]

[0089]

SEQUENCE LISTING

<110> RIKEN

<120> Dye and fluorescent proteins

<130> A31700A

<160> 16

<210> 1

<211> 221

<212> PRT

<213> *Montipora. sp*

<400> 1

Met Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly

1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys

20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly

35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Leu Ser Gln Tyr Gly

50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys

65 70 75 80

Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met His Phe Glu

85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn

100 105 110

Cys Phe Ile Tyr Asn Val Lys Ile Ser Gly Val Asn Phe Pro Pro Asn

115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu

130 135 140

Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala

145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr

165 170 175

Tyr Lys Ala Lys Lys Pro Val Arg Met Pro Gly Tyr His Tyr Val Asp

180 185 190

Arg Lys Leu Asp Val Thr Ser His Asn Lys Asp Tyr Thr Phe Val Glu

195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg His Ser Leu Leu Gly

210 215 220

<210> 2

<211> 666

<212> DNA

<213> *Montipora. sp*

<400> 2

atg agt gtg atc gct aaa caa atg acc tac aag gtt tat atg tca ggc

Met Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly

1 5 10 15

acg gtc aat gga cac tac ttt gag gtc gaa ggc gat gga aaa gga aag

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys

20 25 30

cct tac gag ggg gag cag acg gta aag ctc act gtc acc aag ggt gga

Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly

35 40 45

cct ctg cca ttt gct tgg gat att tta tca cca ctg tct cag tac gga

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Leu Ser Gln Tyr Gly

50 55 60

agc ata cca ttc acc aag tac cct gaa gac atc cct gat tat gta aag

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys

65 70 75 80

cag tca ttc cct gag gga tat aca tgg gag agg atc atg cac ttt gaa

Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met His Phe Glu

85 90 95

gat ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggc aac

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn

100 105 110

tgt ttc atc tac aat gtc aaa atc tct ggt gtg aac ttt cct ccc aat

Cys Phe Ile Tyr Asn Val Lys Ile Ser Gly Val Asn Phe Pro Pro Asn

115 120 125

gga cct gtt atg cag aag aag aca cag ggc tgg gaa ccc aac act gag

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu

130 135 140

cgt ctc ttt gca cga gat gga atg ctg ata gga aac aac ttt atg gct

Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala

145 150 155 160

ctg aag ttg gaa gga ggt ggt cac tat ttg tgt gaa ttc aaa tct act

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr

165 170 175

tac aag gca aag aag cct gtg agg atg cca ggg tat cac tat gtt gac

Tyr Lys Ala Lys Lys Pro Val Arg Met Pro Gly Tyr His Tyr Val Asp

180 185 190

cgc aaa ctg gat gta acc agt cac aac aag gat tac aca ttt gtt gag

Arg Lys Leu Asp Val Thr Ser His Asn Lys Asp Tyr Thr Phe Val Glu

195 200 205

cag tgt gaa ata tcc att gca cgc cac tct ttg ctc ggt tga

Gln Cys Glu Ile Ser Ile Ala Arg His Ser Leu Leu Gly

210 215 220

<210> 3

<211> 222

<212> PRT

<213> *Montipora. sp*

<400> 3

Met Val Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser

1 5 10 15

Gly Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly

20 25 30

Lys Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly

35 40 45

Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Leu Ser Gln Tyr

50 55 60

Gly Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val

65 70 75 80

Lys Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe

85 90 95

Glu Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly

100 105 110

Asn Cys Phe Ile Tyr Asn Val Lys Ile Ser Gly Val Asn Phe Pro Pro

115 120 125

Asn Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Ser Thr

130 135 140

Glu Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asp Phe Met

145 150 155 160

Ala Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser

165 170 175

Thr Tyr Lys Ala Lys Lys Pro Val Arg Met Pro Gly Tyr His Tyr Val

180 185 190

Asp Arg Lys Leu Asp Val Thr Ser His Asn Arg Asp Tyr Thr Ser Val

195 200 205

Glu Gln Cys Glu Ile Ser Ile Ala Arg His Ser Leu Leu Gly

210 215 220

<210> 4

<211> 669

<212> DNA

<213> *Montipora. sp*

<400> 4

atg gtg agt gtg atc gct aaa caa atg acc tac aag gtt tat atg tca

Met Val Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser

1 5 10 15

ggc acg gtc aat gga cac tac ttt gag gtc gaa ggc gat gga aaa gga

Gly Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly

20 25 30

aag cct tac gag gga gag cag aca gta aag ctc act gtc acc aag ggt

Lys Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly

35 40 45

gga cct ctg cca ttt gct tgg gat att tta tca cca ctg tct cag tac

Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Leu Ser Gln Tyr

50 55 60

gga agc ata cca ttc acc aag tac cct gaa gac atc cct gat tat gta

Gly Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val

65 70 75 80

aag cag tca ttc cct gag gga tat aca tgg gag agg atc atg aac ttt

Lys Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe

85 90 95

gaa gat ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggc

Glu Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly

100 105 110

aac tgt ttc atc tac aat gtc aaa atc tct ggt gtg aac ttt cct ccc

Asn Cys Phe Ile Tyr Asn Val Lys Ile Ser Gly Val Asn Phe Pro Pro

115 120 125

aat gga cct gtt atg cag aag aag aca cag ggc tgg gaa ccc agc act

Asn Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Ser Thr

130 135 140

gag cgt ctc ttt gca cga gat gga atg ctg ata gga aac gat ttt atg

Glu Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asp Phe Met

145 150 155 160

gct ctg aag ttg gaa gga ggt ggt cac tat ttg tgt gaa ttc aaa tct

Ala Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser

165 170 175

act tac aag gca aag aag cct gtg agg atg cca ggg tat cac tat gtt

Thr Tyr Lys Ala Lys Lys Pro Val Arg Met Pro Gly Tyr His Tyr Val

180 185 190

gac cgc aaa ctg gat gta acc agt cac aac agg gat tac aca tct gtt

Asp Arg Lys Leu Asp Val Thr Ser His Asn Arg Asp Tyr Thr Ser Val

195 200 205

gag cag tgt gaa ata tcc att gca cgc cac tct ttg ctc ggt tga

Glu Gln Cys Glu Ile Ser Ile Ala Arg His Ser Leu Leu Gly

210 215 220

<210> 5

<211> 222

<212> PRT

<213> *Montipora. sp*

<400> 5

Met Val Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser

1 5 10 15

Gly Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly

20 25 30

Lys Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly

35 40 45

Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Leu Phe Gln Tyr

50 55 60

Gly Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val

65 70 75 80

Lys Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Thr Met Asn Phe

85 90 95

Glu Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly

100 105 110

Asn Cys Phe Ile Tyr Asn Val Lys Ile Ser Gly Thr Asn Phe Pro Pro

115 120 125

Asn Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Ser Thr

130 135 140

Glu Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asp Tyr Met

145 150 155 160

Ala Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser
165 170 175

Thr Tyr Lys Ala Lys Lys Pro Val Arg Met Pro Gly Tyr His Tyr Ile
180 185 190

Asp Arg Lys Leu Asp Val Thr Ser His Asn Arg Asp Tyr Thr Ser Val
195 200 205

Glu Gln Cys Glu Ile Ala Ile Ala Arg His Ser Leu Leu Gly
210 215 220

<210> 6

<211> 669

<212> DNA

<213> *Montipora. sp*

<400> 6

atg gtg agt gtg atc gct aaa caa atg acc tac aag gtt tat atg tca

Met Val Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser
1 5 10 15

ggc acg gtc aat gga cac tac ttt gag gtc gaa ggc gat gga aaa gga

Gly Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly
20 25 30

aag cct tac gag gga gag cag aca gta aag ctc act gtc acc aag ggt

Lys Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly
35 40 45

gga cct ctg cca ttt gct tgg gat att tta tca cca ctg ttt cag tac

Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Leu Phe Gln Tyr
50 55 60

gga agc ata cca ttc acc aag tac cct gaa gac atc cct gat tat gta

Gly Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val
65 70 75 80

aag cag tca ttc cct gag gga tat aca tgg gag agg acc atg aac ttt

Lys Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Thr Met Asn Phe

85 90 95

gaa gat ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggc
Glu Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly
100 105 110

aac tgt ttc atc tac aat gtc aaa atc tct ggt acg aac ttt cct ccc
Asn Cys Phe Ile Tyr Asn Val Lys Ile Ser Gly Thr Asn Phe Pro Pro
115 120 125

aat gga cct gtt atg cag aag aag aca cag ggc tgg gaa ccc agc act
Asn Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Ser Thr
130 135 140

gag cgt ctc ttt gca cga gat gga atg ctg ata gga aac gat tat atg
Glu Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asp Tyr Met
145 150 155 160

gct ctg aag ttg gaa gga ggt ggt cac tat ttg tgt gaa ttt aaa tct
Ala Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser
165 170 175

act tac aag gca aag aag cct gtg agg atg cca ggg tat cac tat att
Thr Tyr Lys Ala Lys Lys Pro Val Arg Met Pro Gly Tyr His Tyr Ile
180 185 190

gac cgc aaa ctg gat gta acc agt cac aac agg gat tac aca tct gtt
Asp Arg Lys Leu Asp Val Thr Ser His Asn Arg Asp Tyr Thr Ser Val
195 200 205

gag cag tgt gaa ata gcc att gca cgc cac tct ttg ctc ggt tga
Glu Gln Cys Glu Ile Ala Ile Ala Arg His Ser Leu Leu Gly
210 215 220

<210> 7

<211> 222

<212> PRT

<213> *Montipora. sp*

<400> 7

Met Val Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser

1   5   10   15

Gly Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly

            20   25   30

Lys Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly

        35   40   45

Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Leu Met Cys Tyr

    50   55   60

Gly Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val

65   70   75   80

Lys Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Thr Met Asn Phe

                85   90   95

Glu Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly

            100   105   110

Asn Cys Phe Ile Tyr Asn Val Lys Ile Ser Gly Thr Asn Phe Pro Pro

        115   120   125

Asn Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Ser Thr

    130   135   140

Glu Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asp Tyr Met

145   150   155   160

Ala Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser

                165   170   175

Thr Tyr Lys Ala Lys Lys Pro Val Arg Met Pro Gly Tyr His Tyr Ile

            180   185   190

Asp Arg Lys Leu Asp Val Thr Ser His Asn Arg Asp Tyr Thr Ser Val

        195   200   205

Glu Gln Cys Glu Ile Ala Ile Ala Arg His Ser Leu Leu Gly

210 215 220

<210> 8

<211> 669

<212> DNA

<213> *Montipora. sp*

<400> 8

atg gtg agt gtg atc gct aaa caa atg acc tac aag gtt tat atg tca

Met Val Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser

1 5 10 15

ggc acg gtc aat gga cac tac ttt gag gtc gaa ggc gat gga aaa gga

Gly Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly

20 25 30

aag cct tac gag gga gag cag aca gta aag ctc act gtc acc aag ggt

Lys Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly

35 40 45

gga cct ctg cca ttt gct tgg gat att tta tca cca ctg atg tgt tac

Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Leu Met Cys Tyr

50 55 60

gga agc ata cca ttc acc aag tac cct gaa gac atc cct gat tat gta

Gly Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val

65 70 75 80

aag cag tca ttc cct gag gga tat aca tgg gag agg acc atg aac ttt

Lys Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Thr Met Asn Phe

85 90 95

gaa gat ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggc

Glu Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly

100 105 110

aac tgt ttc atc tac aat gtc aaa atc tct ggt acg aac ttt cct ccc
Asn Cys Phe Ile Tyr Asn Val Lys Ile Ser Gly Thr Asn Phe Pro Pro
115 120 125

aat gga cct gtt atg cag aag aag aca cag ggc tgg gaa ccc agc act
Asn Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Ser Thr
130 135 140

gag cgt ctc ttt gca cga gat gga atg ctg ata gga aac gat tat atg
Glu Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asp Tyr Met
145 150 155 160

gct ctg aag ttg gaa gga ggt ggt cac tat ttg tgt gaa ttt aaa tct
Ala Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser
165 170 175

act tac aag gca aag aag cct gtg agg atg cca ggg tat cac tat att
Thr Tyr Lys Ala Lys Lys Pro Val Arg Met Pro Gly Tyr His Tyr Ile
180 185 190

gac cgc aaa ctg gat gta acc agt cac aac agg gat tac aca tct gtt
Asp Arg Lys Leu Asp Val Thr Ser His Asn Arg Asp Tyr Thr Ser Val
195 200 205

gag cag tgt gaa ata gcc att gca cgc cac tct ttg ctc ggt tga
Glu Gln Cys Glu Ile Ala Ile Ala Arg His Ser Leu Leu Gly
210 215 220

<210> 9

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 9

gaaggrtgyg tcaayggrca y 21

<210> 10

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 10

acvggdccat ydgvaagaaa rtt 23

<210> 11

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 11

ggccacgcgt cgactagtac gggiigggii gggiig 36

<210> 12

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 12

ctcagggaat gactgcttta cat 23

<210> 13

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 13

ggccacgcgt cgactagtac 20

<210> 14

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 14

gtcttcaggg tacttggtga 20

<210> 15

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 15

atgtaaagca gtcattccct gag 23

<210> 16

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

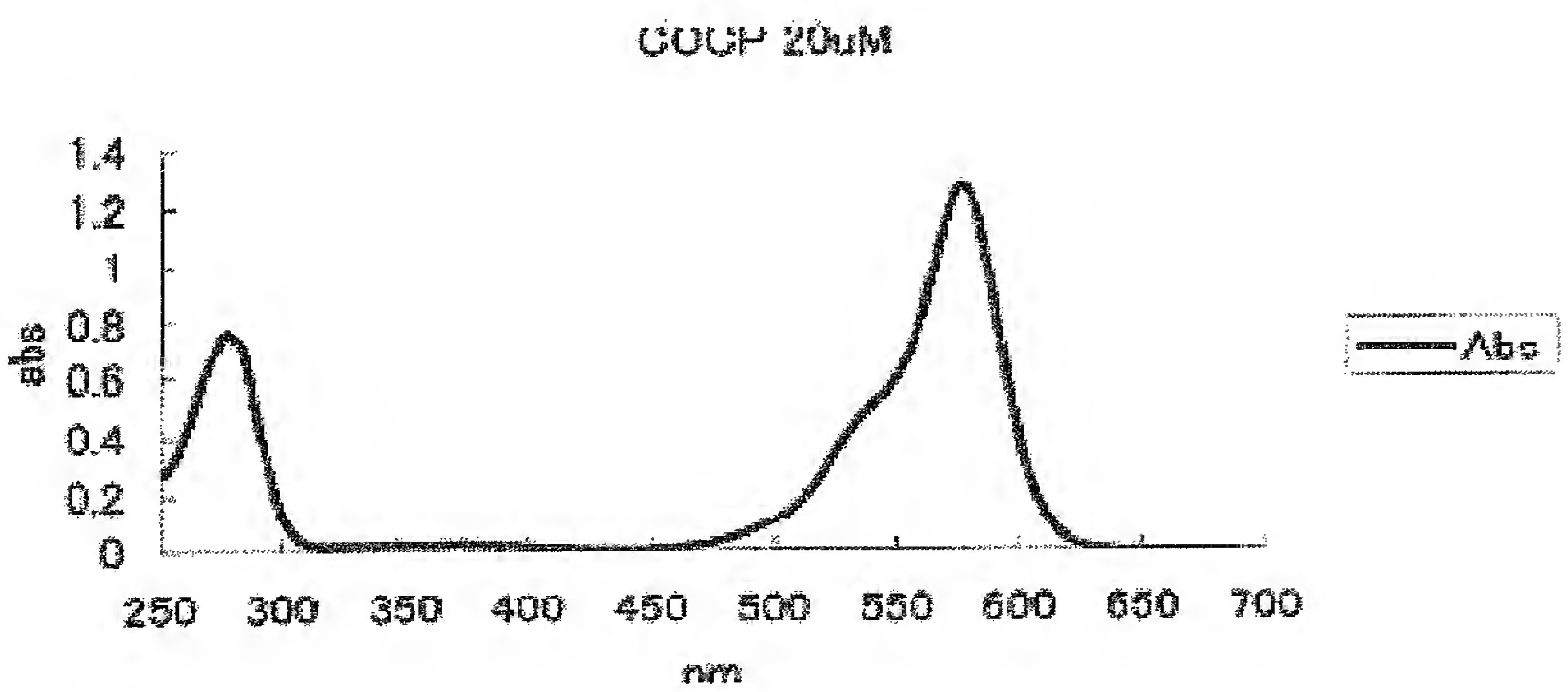
<223> Description of Artificial Sequence: Synthetic DNA

<400> 16

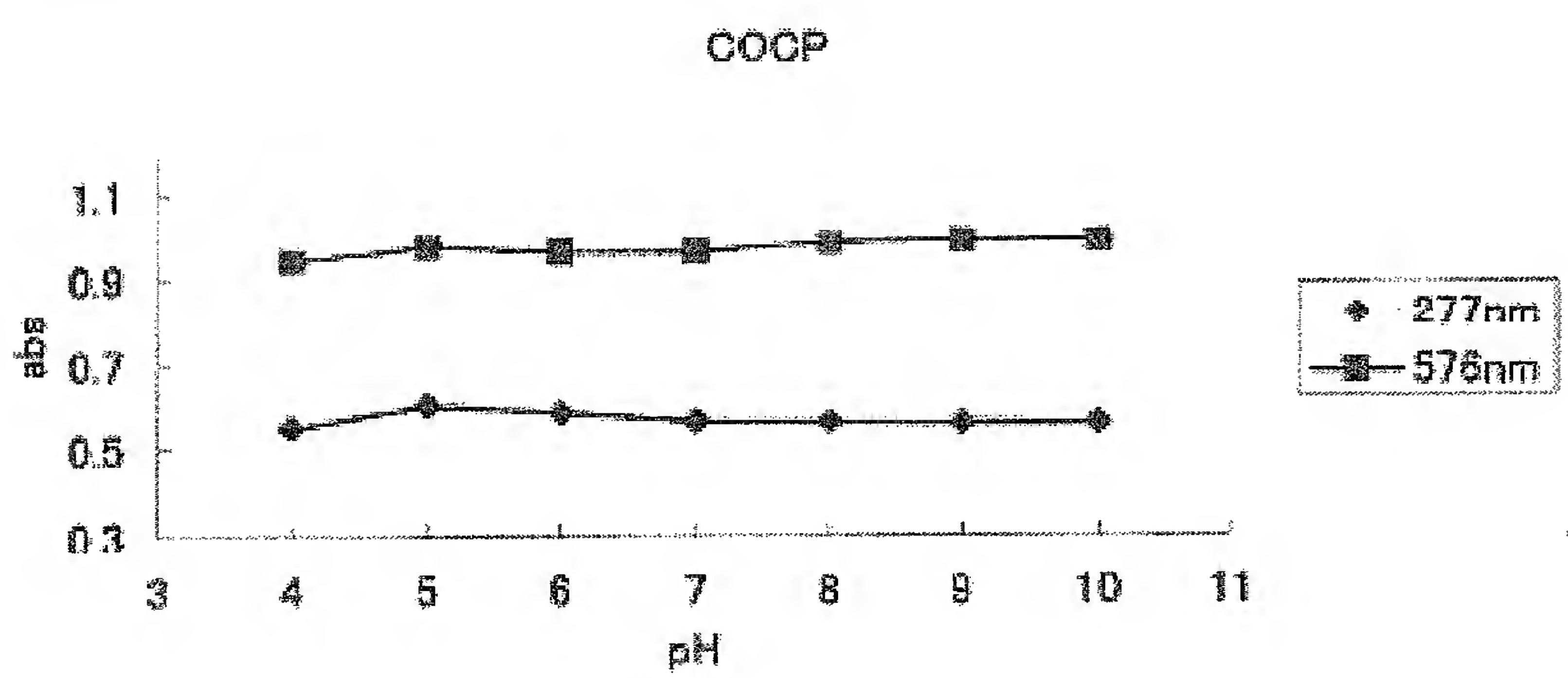
cccggatccg accatggcta ccttggttaa aga 33

[Name of Document] Drawings

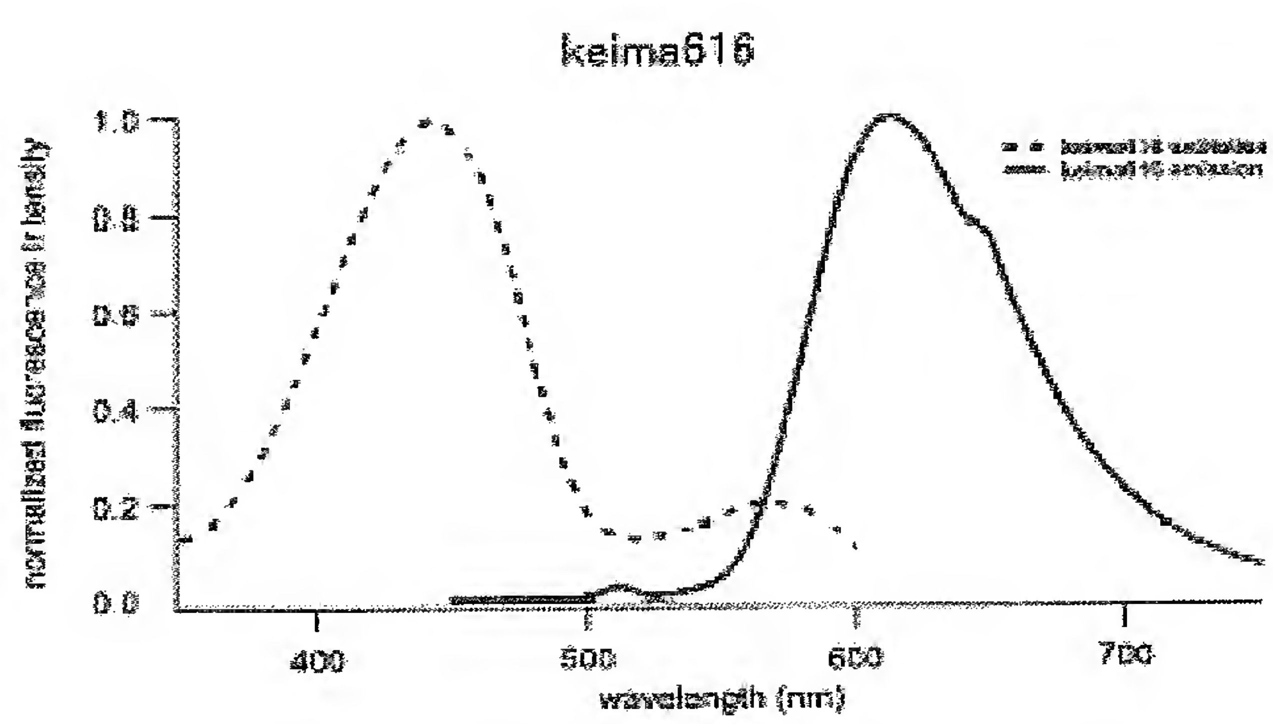
[Figure 1]



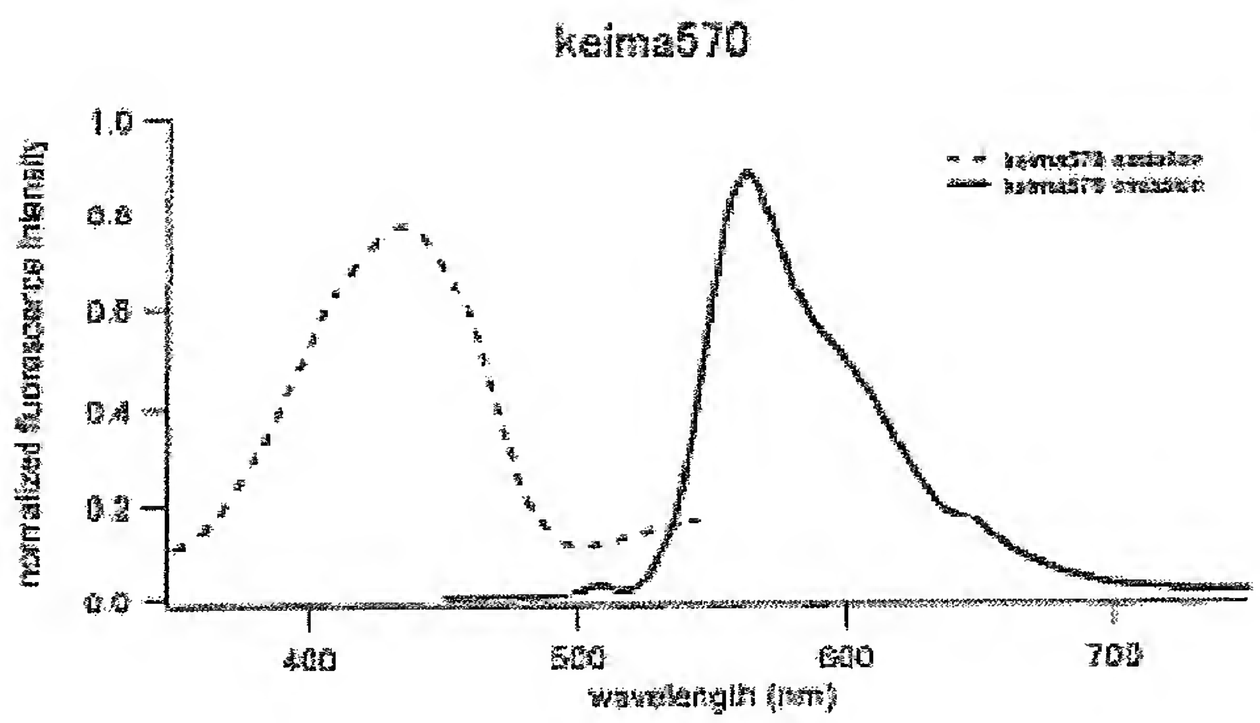
[Figure 2]



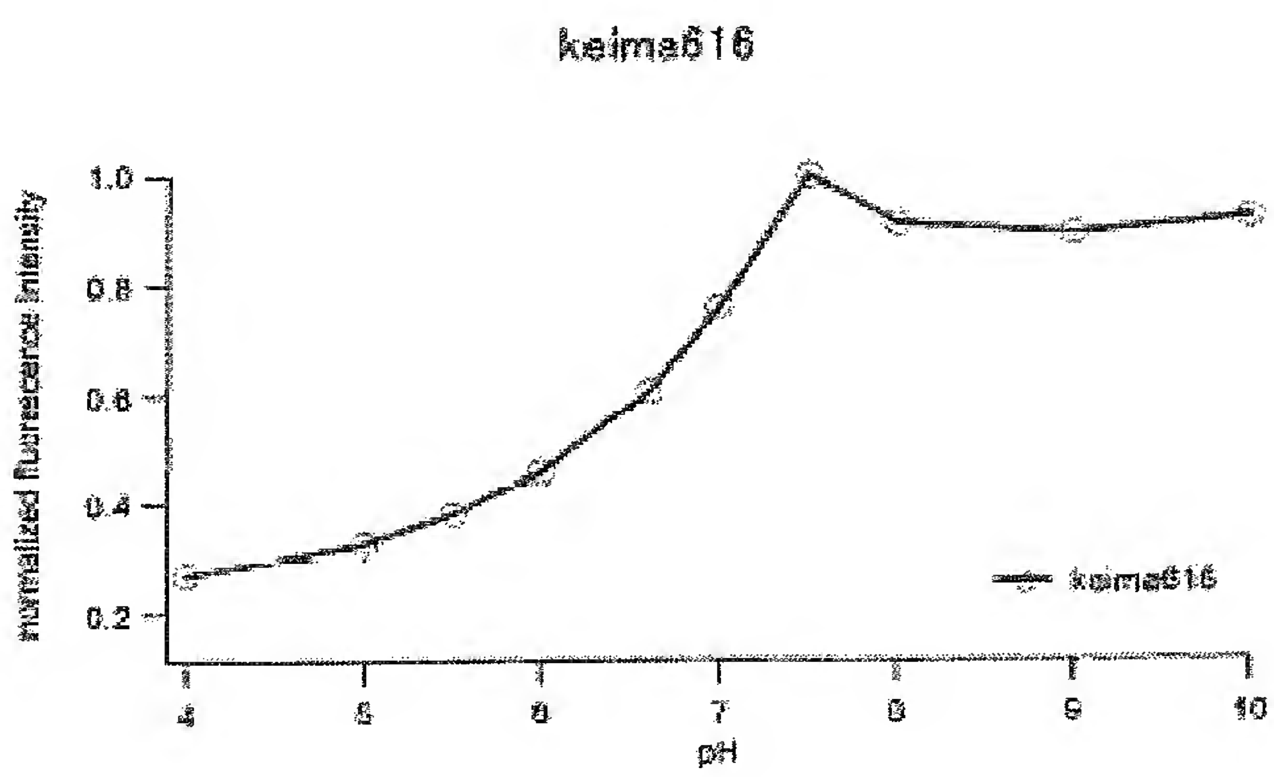
[Figure 3]



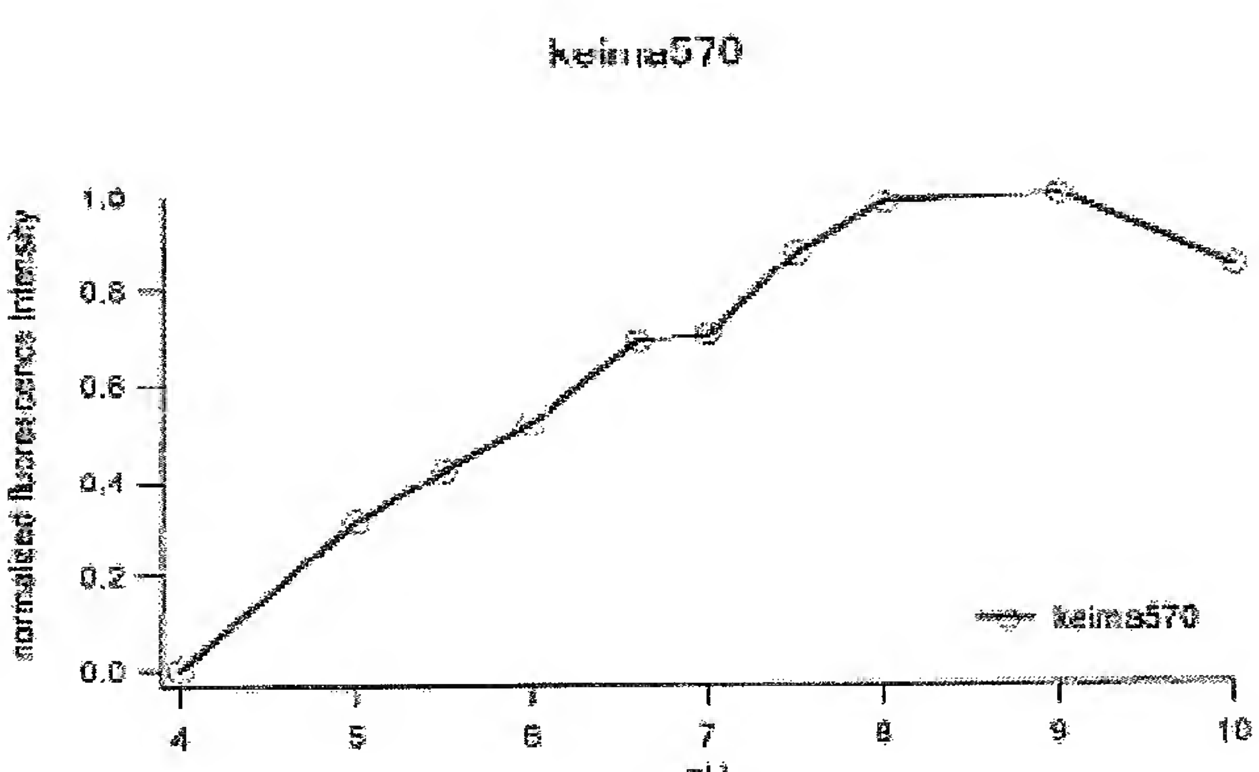
[Figure 4]



[Figure 5]



[Figure 6]



[Name of Document] ABSTRACT

[Abstract]

[Object] Tto provide a red or orange fluorescent protein, which is characterized in that the difference (stokes shift) between an excitation peak value (wavelength of maximum absorption) and a fluorescence peak value (wavelength of maximum fluorescence) is greatened, so that the maximum fluorescence can be obtained by the maximum excitation.

[Means for solution] A fluorescent protein described in the following (a) or (b):

(a) a protein having the amino acid sequence shown in SEQ ID NO: 5 or 7; or

(b) a protein, which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 5 or 7, which has fluorescence properties, and which has a stokes shift of 100 nm or greater.

[Selected Drawing] None